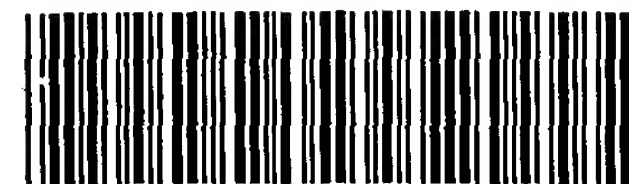


0590
1/31

OIPE

RAW SEQUENCE LISTING

DATE: 02/11/2002

PATENT APPLICATION: US/10/021,758

TIME: 09:49:15

Input Set : N:\Crf3\RULE60\10021758.txt

Output Set: N:\CRF3\02112002\J021758.raw

ENTERED

```

4 <110> APPLICANT: Conklin, Darrell C.
5      Yamamoto, Gayle
6      Gao, Zeren
7      Whitmore, Theodore E.
8      Jaspers, Stephen
10 <120> TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING
11      BETA-1,3-GALACTOSYLTRANSFERASE HOMOLOG, ZNSSP6
13 <130> FILE REFERENCE: 98-80
15 <140> CURRENT APPLICATION NUMBER: 10/021,758
16 <141> CURRENT FILING DATE: 2001-10-22
18 <150> PRIOR APPLICATION NUMBER: 09/482,180
19 <151> PRIOR FILING DATE: 2000-01-12
21 <150> PRIOR APPLICATION NUMBER: 60/115,721
22 <151> PRIOR FILING DATE: 1999-01-12
24 <160> NUMBER OF SEQ ID NOS: 20
26 <170> SOFTWARE: FastSEQ for Windows Version 3.0
28 <210> SEQ ID NO: 1
29 <211> LENGTH: 1420
30 <212> TYPE: DNA
31 <213> ORGANISM: Homo sapiens
33 <220> FEATURE:
34 <221> NAME/KEY: CDS
35 <222> LOCATION: (135)...(1271)
37 <400> SEQUENCE: 1
38  cacgagctcc acgcccgtac cccggcggtca cgctcagccc gcggtgctcg cacacctgag      60
39  actcatctcg ctctgacccc gccgccgccg ccgcccggca tcctgagcac ggagacagtc      120
40  tccagctgcc gtte atg ctt cct ccc cag cct tcc gca gcc cac cag gga      170
41      Met Leu Pro Pro Gln Pro Ser Ala Ala His Gln Gly
42      1              5              10
44  agg ggc ggt agg agt ggc ctt tta cca aag gga ccg gcg atg ctc tgc      218
45  Arg Gly Gly Arg Ser Gly Leu Leu Pro Lys Gly Pro Ala Met Leu Cys
46      15              20              25
48  agg ctg tgc tgg ctg gtc tgc tac agc ttg gct gtg ctg ttg ctc ggc      266
49  Arg Leu Cys Trp Leu Val Ser Tyr Ser Leu Ala Val Leu Leu Leu Gly
50      30              35              40
52  tgc ctg ctc ttc ctg agg aag gcg gcc aag ccc gca gga gac ccc acg      314
53  Cys Leu Leu Phe Leu Arg Lys Ala Ala Lys Pro Ala Gly Asp Pro Thr
54      45              50              55              60
56  gcc cac cag cct ttc tgg gct ccc cca aca ccc cgt cac agc cgg tgt      362
57  Ala His Gln Pro Phe Trp Ala Pro Pro Thr Pro Arg His Ser Arg Cys
58      65              70              75
60  cca ccc aac cac aca gtg tct agc gcc tct ctg tcc ctg cct agc cgt      410
61  Pro Pro Asn His Thr Val Ser Ser Ala Ser Leu Ser Leu Pro Ser Arg

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/021,758

DATE: 02/11/2002

TIME: 09:49:15

Input Set : N:\Crf3\RULE60\10021758.txt

Output Set: N:\CRF3\02112002\J021758.raw

62		80		85		90		
64	cac cgt ctc ttc ttg acc tat cgt cac tgc cga aat ttc tct atc ttg							458
65	His Arg Leu Phe Leu Thr Tyr Arg His Cys Arg Asn Phe Ser Ile Leu							
66		95		100		105		
68	ctg gag cct tca ggc tgt tcc aag gat acc ttc ttg ctc ctg gcc atc							506
69	Leu Glu Pro Ser Gly Cys Ser Lys Asp Thr Phe Leu Leu Leu Ala Ile							
70		110		115		120		
72	aag tca cag cct ggt cac gtg gag cga cgt gcg gct atc cgc agc acg							554
73	Lys Ser Gln Pro Gly His Val Glu Arg Arg Ala Ala Ile Arg Ser Thr							
74	125		130		135		140	
76	tgg ggc agg gtg ggg gga tgg gct agg ggc cgg cag ctg aag ctg gtg							602
77	Trp Gly Arg Val Gly Gly Trp Ala Arg Gly Arg Gln Leu Lys Leu Val							
78		145		150		155		
80	ttc ctc cta ggg gtg gca gga tcc gct ccc cca gcc cag ctg ctg gcc							650
81	Phe Leu Leu Gly Val Ala Gly Ser Ala Pro Pro Ala Gln Leu Leu Ala							
82		160		165		170		
84	tat gag agt agg gag ttt gat gac atc ctc cag tgg gac ttc act gag							698
85	Tyr Glu Ser Arg Glu Phe Asp Asp Ile Leu Gln Trp Asp Phe Thr Glu							
86		175		180		185		
88	gac ttc ttc aac ctg acg ctc aag gag ctg cac ctg cag cgc tgg gtg							746
89	Asp Phe Phe Asn Leu Thr Leu Lys Glu Leu His Leu Gln Arg Trp Val							
90		190		195		200		
92	gtg gct gcc tgc ccc cag gcc cat ttc atg cta aag gga gat gac gat							794
93	Val Ala Ala Cys Pro Gln Ala His Phe Met Leu Lys Gly Asp Asp Asp							
94	205		210		215		220	
96	gtc ttt gtc cac gtc ccc aac gtg tta gag ttc ctg gat ggc tgg gac							842
97	Val Phe Val His Val Pro Asn Val Leu Glu Phe Leu Asp Gly Trp Asp							
98		225		230		235		
100	cca gcc cag gac ctc ctg gtg gga gat gtc atc cgc caa gcc ctg ccc							890
101	Pro Ala Gln Asp Leu Leu Val Gly Asp Val Ile Arg Gln Ala Leu Pro							
102		240		245		250		
104	aac agg aac act aag gtc aaa tac ttc atc cca ccc tca atg tac agg							938
105	Asn Arg Asn Thr Lys Val Lys Tyr Phe Ile Pro Pro Ser Met Tyr Arg							
106		255		260		265		
108	gcc acc cac tac cca ccc tat gct ggt ggg gga gga tat gtc atg tcc							986
109	Ala Thr His Tyr Pro Pro Tyr Ala Gly Gly Gly Gly Tyr Val Met Ser							
110		270		275		280		
112	aga gcc aca gtg cgg cgc ctc cag gct atc atg gaa gat gct gaa ctc							1034
113	Arg Ala Thr Val Arg Arg Leu Gln Ala Ile Met Glu Asp Ala Glu Leu							
114	285		290		295		300	
116	ctc tcc att gat gat gtc ttt gtg ggt atg tgc ctg agg agg ctg ggg							1082
117	Leu Ser Ile Asp Asp Val Phe Val Gly Met Cys Leu Arg Arg Leu Gly							
118		305		310		315		
120	ctg agc cct atg cac cat gct ggc ttc aag aca ttt gga atc cgg cgg							1130
121	Leu Ser Pro Met His His Ala Gly Phe Lys Thr Phe Gly Ile Arg Arg							
122		320		325		330		
124	ccc ctg gac ccc tta gac ccc tgc ctg tat agg ggg ctc ctg ctg gtt							1178
125	Pro Leu Asp Pro Leu Asp Pro Cys Leu Tyr Arg Gly Leu Leu Leu Val							
126		335		340		345		

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/021,758

DATE: 02/11/2002

TIME: 09:49:15

Input Set : N:\Crf3\RULE60\10021758.txt

Output Set: N:\CRF3\02112002\J021758.raw

```

128  cac cgc ctc agc ccc ctc gag atg tgg acc atg tgg gca ctg gtg aca      1226
129  His Arg Leu Ser Pro Leu Glu Met Trp Thr Met Trp Ala Leu Val Thr
130      350                      355                      360
132  gat gag ggg ctc aag tgt gca gct ggc ccc ata ccc cag cgc tga      1271
133  Asp Glu Gly Leu Lys Cys Ala Ala Gly Pro Ile Pro Gln Arg  *
134  365                      370                      375
136  aggggtgggtt gggcaacagc ctgagagtgg actcagtgtt gattctctat cgtgatgcga      1331
137  aattgatgcc tgctgctcta cagaaaatgc caacttggtt ttttaactcc tctcaccctg      1391
138  ttagctctga ttaaaaacac tgcaaccca      1420
140 <210> SEQ ID NO: 2
141 <211> LENGTH: 378
142 <212> TYPE: PRT
143 <213> ORGANISM: Homo sapiens
145 <400> SEQUENCE: 2
146  Met Leu Pro Pro Gln Pro Ser Ala Ala His Gln Gly Arg Gly Gly Arg
147      1                      5                      10                      15
148  Ser Gly Leu Leu Pro Lys Gly Pro Ala Met Leu Cys Arg Leu Cys Trp
149      20                      25                      30
150  Leu Val Ser Tyr Ser Leu Ala Val Leu Leu Leu Gly Cys Leu Leu Phe
151      35                      40                      45
152  Leu Arg Lys Ala Ala Lys Pro Ala Gly Asp Pro Thr Ala His Gln Pro
153      50                      55                      60
154  Phe Trp Ala Pro Pro Thr Pro Arg His Ser Arg Cys Pro Pro Asn His
155      65                      70                      75                      80
156  Thr Val Ser Ser Ala Ser Leu Ser Leu Pro Ser Arg His Arg Leu Phe
157      85                      90                      95
158  Leu Thr Tyr Arg His Cys Arg Asn Phe Ser Ile Leu Leu Glu Pro Ser
159      100                     105                     110
160  Gly Cys Ser Lys Asp Thr Phe Leu Leu Leu Ala Ile Lys Ser Gln Pro
161      115                     120                     125
162  Gly His Val Glu Arg Arg Ala Ala Ile Arg Ser Thr Trp Gly Arg Val
163      130                     135                     140
164  Gly Gly Trp Ala Arg Gly Arg Gln Leu Lys Leu Val Phe Leu Leu Gly
165      145                     150                     155                     160
166  Val Ala Gly Ser Ala Pro Pro Ala Gln Leu Leu Ala Tyr Glu Ser Arg
167      165                     170                     175
168  Glu Phe Asp Asp Ile Leu Gln Trp Asp Phe Thr Glu Asp Phe Phe Asn
169      180                     185                     190
170  Leu Thr Leu Lys Glu Leu His Leu Gln Arg Trp Val Val Ala Ala Cys
171      195                     200                     205
172  Pro Gln Ala His Phe Met Leu Lys Gly Asp Asp Asp Val Phe Val His
173      210                     215                     220
174  Val Pro Asn Val Leu Glu Phe Leu Asp Gly Trp Asp Pro Ala Gln Asp
175      225                     230                     235                     240
176  Leu Leu Val Gly Asp Val Ile Arg Gln Ala Leu Pro Asn Arg Asn Thr
177      245                     250                     255
178  Lys Val Lys Tyr Phe Ile Pro Pro Ser Met Tyr Arg Ala Thr His Tyr
179      260                     265                     270
180  Pro Pro Tyr Ala Gly Gly Gly Gly Tyr Val Met Ser Arg Ala Thr Val

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/021,758

DATE: 02/11/2002

TIME: 09:49:15

Input Set : N:\Crf3\RULE60\10021758.txt

Output Set: N:\CRF3\02112002\J021758.raw

```

181          275          280          285
182 Arg Arg Leu Gln Ala Ile Met Glu Asp Ala Glu Leu Leu Ser Ile Asp
183          290          295          300
184 Asp Val Phe Val Gly Met Cys Leu Arg Arg Leu Gly Leu Ser Pro Met
185          305          310          315          320
186 His His Ala Gly Phe Lys Thr Phe Gly Ile Arg Arg Pro Leu Asp Pro
187          325          330          335
188 Leu Asp Pro Cys Leu Tyr Arg Gly Leu Leu Leu Val His Arg Leu Ser
189          340          345          350
190 Pro Leu Glu Met Trp Thr Met Trp Ala Leu Val Thr Asp Glu Gly Leu
191          355          360          365
192 Lys Cys Ala Ala Gly Pro Ile Pro Gln Arg
193          370          375

```

```

195 <210> SEQ ID NO: 3
196 <211> LENGTH: 1134
197 <212> TYPE: DNA
198 <213> ORGANISM: Artificial Sequence
200 <220> FEATURE:
201 <223> OTHER INFORMATION: degenerate sequence
203 <221> NAME/KEY: variation
204 <222> LOCATION: (1)...(1134)
205 <223> OTHER INFORMATION: n is any nucleotide
207 <221> NAME/KEY: misc_feature
208 <222> LOCATION: (1)...(1134)
209 <223> OTHER INFORMATION: n = A,T,C or G
211 <400> SEQUENCE: 3

```

```

W--> 212 atgytncnc cncarccnws ngcngcncay carggnmgng gnggnmgngws nggnytnytn      60
W--> 213 ccnaarggnc cngcnatgyt ntgymgnytn tgytggytng tnwsntayws nytngcngtn      120
W--> 214 ytnytnytn gntgyytnyt nttyytnmgn aargcngcna arccngcngg ngayccnacd      180
W--> 215 gencaycarc cnttytgggc nccnccnacd ccnmgncayw snmgntgycc nccnaaycay      240
W--> 216 acngtnwsnw sngcnwsnyt nwsnytnccn wsnmgncaym gnytnnttyt nacntaymgn      300
W--> 217 caytgymgna aytywsnat hytnytngar ccnwsnggnt gywsnaarga yacnttyytn      360
W--> 218 ytnytngcna thaarwsnca rccnggncay gtngarmgmn gngcngcnat hmgwnsnacd      420
W--> 219 tggggngmgng tngggngntg ggcnmgnggn mgncarytna arytngtntt yytnytnngn      480
W--> 220 gtngcnggnw sngcncncnc ngcncarytn ytngcntayg arwsnmngna rttygaygay      540
W--> 221 athytncart gggayttyac ngargaytty ttyaayytna cnytnaarga rytncayytn      600
W--> 222 carmgntggg tngtngcngc ntgyccncar gcncayttya tgytnaargg ngaygaygay      660
W--> 223 gtnttygtnc aygtncncaa ygtnytngar ttyytngayg gntgggaycc ngcncargay      720
W--> 224 ytnytngtng gngaygtnat hmgncargcn ytncncaaym gnaayacnaa rgtnaartay      780
W--> 225 ttyathccnc cnwsnatgta ymgngcnacd caytayccnc cntaygcngg nggnggnggn      840
W--> 226 taygtnatgw snmgngcnac ngtnmgngmn ytncargcna thatggarga ygcngarytn      900
W--> 227 ytnwsnathg aygaygtntt ygtnggnatg tgyytnmgmn gnytnngnytn nwsnccnatg      960
W--> 228 caycaygcng gnttyaarac nttygggnath mgnmgnccny tngayccnytn ngayccntgy      1020
W--> 229 ytnaymgng gnytnytnyt ngtncaymgn ytnwsnccny tngaratgtg gacnatgtgg      1080
W--> 230 gcnytngtna cngaygargg nytnaartgy gcngcnggnc cnathccnca rmgn      1134

```

```

232 <210> SEQ ID NO: 4
233 <211> LENGTH: 27
234 <212> TYPE: DNA
235 <213> ORGANISM: Artificial Sequence

```


RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/021,758

DATE: 02/11/2002

TIME: 09:49:15

Input Set : N:\Crf3\RULE60\10021758.txt

Output Set: N:\CRF3\02112002\J021758.raw

237 <220> FEATURE:
 238 <223> OTHER INFORMATION: oligonucleotide primer
 240 <400> SEQUENCE: 4
 241 ccatacctaatacgaactcact atagggc 27
 243 <210> SEQ ID NO: 5
 244 <211> LENGTH: 23
 245 <212> TYPE: DNA
 246 <213> ORGANISM: Artificial Sequence
 248 <220> FEATURE:
 249 <223> OTHER INFORMATION: oligonucleotide primer
 251 <400> SEQUENCE: 5
 252 cggatagccg cacgtcgtc cac 23
 254 <210> SEQ ID NO: 6
 255 <211> LENGTH: 23
 256 <212> TYPE: DNA
 257 <213> ORGANISM: Artificial Sequence
 259 <220> FEATURE:
 260 <223> OTHER INFORMATION: oligonucleotide primer
 262 <400> SEQUENCE: 6
 263 actcactata gggctcgagc ggc 23
 265 <210> SEQ ID NO: 7
 266 <211> LENGTH: 23
 267 <212> TYPE: DNA
 268 <213> ORGANISM: Artificial Sequence
 270 <220> FEATURE:
 271 <223> OTHER INFORMATION: oligonucleotide primer
 273 <400> SEQUENCE: 7
 274 tgaccaggct gtgacttgat ggc 23
 276 <210> SEQ ID NO: 8
 277 <211> LENGTH: 24
 278 <212> TYPE: DNA
 279 <213> ORGANISM: Artificial Sequence
 281 <220> FEATURE:
 282 <223> OTHER INFORMATION: oligonucleotide primer
 284 <400> SEQUENCE: 8
 285 cttggcacga ggcacgagct ccac 24
 287 <210> SEQ ID NO: 9
 288 <211> LENGTH: 24
 289 <212> TYPE: DNA
 290 <213> ORGANISM: Artificial Sequence
 292 <220> FEATURE:
 293 <223> OTHER INFORMATION: oligonucleotide primer
 295 <400> SEQUENCE: 9
 296 ctacaggctgt tgcccaaccc accc 24
 298 <210> SEQ ID NO: 10
 299 <211> LENGTH: 25
 300 <212> TYPE: DNA
 301 <213> ORGANISM: Artificial Sequence
 303 <220> FEATURE:

VERIFICATION SUMMARY

PATENT APPLICATION: US/10/021,758

DATE: 02/11/2002

TIME: 09:49:16

Input Set : N:\Crf3\RULE60\10021758.txt

Output Set: N:\CRF3\02112002\J021758.raw

L:212 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3
L:213 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3
L:214 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3
L:215 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3
L:216 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3
L:217 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3
L:218 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3
L:219 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3
L:220 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3
L:221 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3
L:222 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3
L:223 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3
L:224 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3
L:225 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3
L:226 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3
L:227 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3
L:228 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3
L:229 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3
L:230 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3
L:357 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:15
L:440 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:20